EXHIBIT A

Controlled Site-Selective Glycosylation of Proteins by a Combined Site-Directed Mutagenesis and Chemical Modification Approach

Benjamin G. Davis, Richard C. Lloyd, and J. Bryan Jones*

Department of Chemistry, University of Toronto, 80 St. George Street, Toronto, Ontario M5S 3H6, Canada

Received August 17, 1998

Surface glycoproteins act as markers in cell–cell communication events that determine microbial virulence,[1] inflammation[2] and host immune responses.[3] In addition, the correct glycosylation of proteins is critical to their expression and folding[4] and increases their thermal and proteolytic stability.[5] Glycoproteins occur naturally in a number of forms (glycoforms)[6] that possess the same peptide backbone but differ in both the nature and site of glycosylation. The differences exhibited[6,7] by each component within these microheterogeneous mixtures present regulatory difficulties[8] and problems in determining exact function. To explore these key properties, there is a pressing need for methods that will not only allow the preparation of pure glycosylated proteins but will also allow the preparation of nonnatural variants for the determination of structure–activity relationships (SARs). The few studies that have compared single glycoforms successfully have required abundant sources and extensive chromatographic separation.[9] Neoglycoproteins,[10] formed via unnatural linkages between sugars and proteins, provide an invaluable alternative source of carbohydrate–protein conjugates.[11] In particular, chemical glycosylation allows control of the glycan structure and the nature of the sugar–protein bond. However, despite these advantages, existing methods for their preparation[11a] typically generate mixtures. Advances in the site-specific glycosylation of bovine serum albumin (BSA) have been made.[12] However, these methods rely upon modification of an existing cysteine in BSA and, as such, allow no flexibility in the choice of glycosylation site. We therefore set ourselves the goal of developing a versatile route to neoglycoproteins without these limitations.

Site-directed mutagenesis combined with chemical modification has permitted us to realize this goal and, for the first time,[13] provides a general method that allows both regio- and glycan-specific glycosylation of proteins. This method is rapid, utilizes reagents that may be prepared in a facile manner and, in principle, is unconstrained in the scope of sites and glycans that may be conjugated. The strategy involves the introduction of cysteine at preselected positions and then reaction of its thiol residue with glycomethanethiosulfonate reagents (Scheme 1). Methanethiosulfonate (MTS) reagents react specifically and quantitatively with thiols[14] and allow the controlled formation of neutral disulfide linkages. Recently, we have successfully used the representative serine protease subtilisin *Bacillus lentus* (SBL) as our vehicle and continue to do so here.[15] SBL is an ideal model protein for evaluating the validity of this strategy as it does not contain a natural cysteine and is not naturally glycosylated.

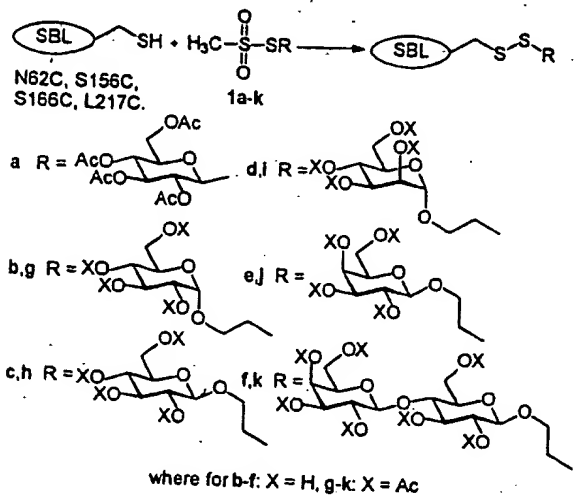
Four SBL sites at different locations and of different characteristics were selected for mutation to cysteine in order to provide a broad test of the glycosylation methodology. S156 of the S_1-pocket is a surface-exposed residue that permits the introduction of externally disposed glycans mirroring those found naturally in glycoproteins.[16] In contrast, N62 in the S_2 pocket, S166 in the S_1 pocket, and L217 in the S_1' pocket have side chains that are internally oriented and test the applicability of the method for introducing sugars at hindered locations. Broad applicability with respect to the sugar moiety was evaluated by using the representative series of protected and deprotected mono- and disaccharide methanethiosulfonates **1a–k**. These were prepared from their parent carbohydrates in good to excellent yields (Schemes 2 and 3). Two types of glycosylating reagents, the anomeric methanethiosulfonate **1a** and the ethyl-tethered methanethiosulfonates **1b,c,g,h**, were prepared from D-glucose (**2a**, Scheme 2). The preparation of these reagents in fully protected **1a,g,h** and deprotected **1b,c** forms allowed the effects of increased steric bulk and hydrophobicity to be assessed. Parallel routes allowed similarly efficient access to the α-D-manno-MTS reagents **1d** and **1i**, which are epimeric at C-2 relative to **1b** and **1g**, respectively, and the β-D-galacto-MTS reagents **1e** and **1j**, epimeric at C-4 relative to **1c** and **1h**, respectively (Scheme 3).

The glyco-MTS reagents **1a–k** were then reacted with SBL-N62C, -S156C, -S166C, and -L217C in aqueous buffer.[15] These reactions were rapid and quantitative, as judged by monitoring changes in specific activity and by titration of residual free thiols with Ellman's reagent.[17] The glycosylated chemically modified mutants (CMMs) were purified by size-exclusion chromatography and dialysis, and their structures were confirmed by rigorous ES–MS analysis. The CMMs each appeared as a single band on nondenaturing gradient PAGE, thereby establishing their high purities. In all cases, modification with the fully deprotected reagents **1b–f** led to site-specific glycosylations and the formation of single glycoforms. These are the first examples of homogeneous neoglycoproteins in which both the site of glycosylation and the strucure of the glycan introduced have been predeter-

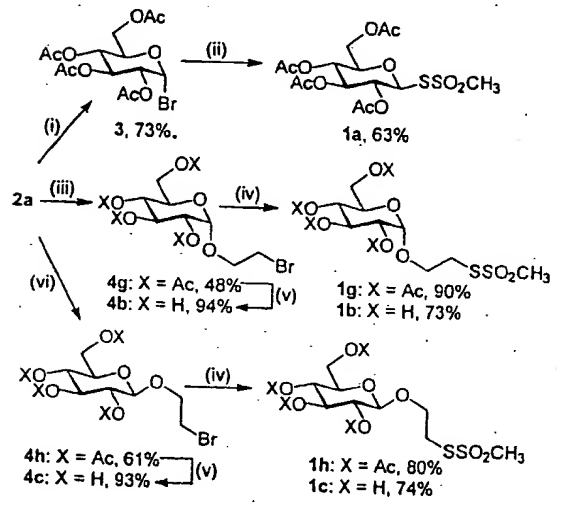
(1) Sharon, N.; Lis, H. *Essays Biochem.* **1995**, *30*, 59–75.
(2) (a) Lasky, L. A. *Annu. Rev. Biochem.* **1995**, *64*, 113–139. (b) Weis, W. I.; Drickamer, K. *Annu. Rev. Biochem.* **1996**, *65*, 441–473.
(3) (a) Varki, A. *Glycobiology* **1993**, *3*, 97–130. (b) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683–720.
(4) Helenius, A. *Mol. Biol. Cell* **1994**, *5*, 253–265.
(5) Opdenakker, G.; Rudd, P. M.; Ponting, C. P.; Dwek, R. A. *FASEB J.* **1993**, *7*, 1330–1337.
(6) Rademacher, T. W.; Parekh, R. B.; Dwek, R. A. *Glycobiology* **1988**, *57*, 785–838.
(7) (a) Parekh, R. B.; Tiemeier, D. C. *Biochemistry* **1989**, *28*, 7670–7679. (b) Knight, P. *Biotechnology* **1989**, *7*, 35–40.
(8) (a) Liu, D. T. Y. *Trends Biotechnol.* **1992**, *10*, 114–120. (b) Bill, R. M.; Flitsch, S. F. *Chem. Biol.* **1996**, *3*, 145–149.
(9) Rudd, P. M.; Dwek, R. A. *Biochemistry* **1994**, *33*, 17–22.
(10) Krantz, M. J.; Holtzmann, N. A.; Stowell, C. P.; Lee, Y. C. *Biochemistry* **1976**, *15*, 3963–3968.
(11) For reviews, see: (a) Stowell, C. P.; Lee, Y. C. *Adv. Carbohydr. Chem. Biochem.* **1980**, *37*, 225–281. (b) *Neoglycoconjugates: Preparation and Applications*; Lee, Y. C., Lee, R. T., Eds.; Academic: London, 1994: (c) Abelson, J. N., Simon., M. I., Eds. *Methods Enzymol.* **1994**, *242*. (d) Lee, Y. C.; Lee, R. T., Eds. *Ibid.* **1994**, *247*.
(12) (a) Davis, N. J.; Flitsch, S. L. *Tetrahedron Lett.* **1991**, *32*, 6793–6796. (b) Wong, S. Y. C.; Guile, G. R.; Dwek, R. A.; Arsequell, G. *Biochem. J.* **1994**, *300*, 843–850. (c) Macindoe, W. M.; van Oijen, A. H.; Boons, G.-J. *J. Chem. Soc., Chem. Commun.* **1998**, 847–848.
(13) During the course of this work, Boons announced his intention to use a similar method for the regioselective glycosylation of IgG (ref 12c) but has published no details.
(14) For reviews see (a) Kenyon, G. L.; Bruice, T. W. *Methods Enzymol.* **1977**, *47*, 407–430. (b) Wynn, R.; Richards, F. M. *Methods Enzymol.* **1995**, *251*, 351–356.
(15) (a) Stabile, M. R.; Lal, W. G.; DeSantis, G.; Gold, M.; Jones, J. B.; Mitchinson, C.; Bott, R. R.; Graycar, T. P.; Liu, C.-C. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2501–2512. (b) Berglund, P.; DeSantis, G.; Stabile, M. R.; Shang, X.; Gold, M.; Bott, R. R.; Graycar, T. P.; Lau, T. H.; Mitchinson, C.; Jones, J. B. *J. Am. Chem. Soc.* **1997**, *119*, 5265–5266.
(16) *Molecular Glycobiology*; Fukuda, M., Hindsgaul, O., Eds.; Oxford University: Oxford, 1994.
(17) Ellman, G. L.; Courtney, K. D.; Andres, V.; Featherstone, R. M. *Biochem. Pharmacol.* **1961**, *7*, 88–95.

10.1021/jo9816461 CCC: $15.00
Published on Web 12/01/1998

Scheme 1



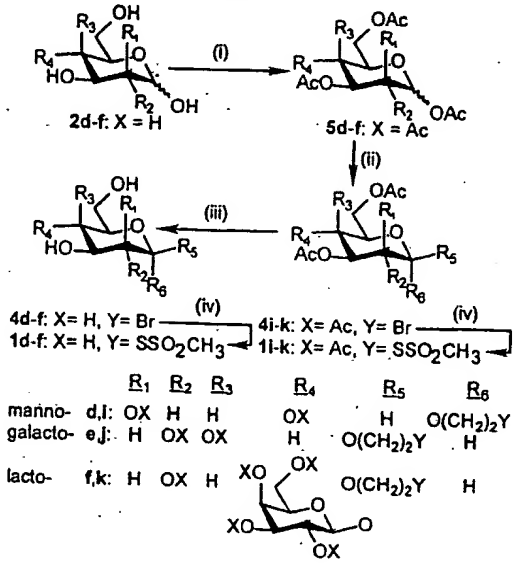
Scheme 2[a]



[a] Reagents and conditions: (i) Ac_2O, py then HBr, AcOH; (ii) $NaSSO_2CH_3$, EtOH, 90 °C; (iii) $Br(CH_2)_2OH$, $BF_3{\cdot}Et_2O$ then Ac_2O, py; (iv) $NaSSO_2CH_3$, DMF, 50 °C; (v) NaOMe, MeOH; (vi) Ac_2O, py then $Br(CH_2)_2OH$, $BF_3{\cdot}Et_2O$, DCM.

mined. Furthermore, modification with the protected MTS **1a,g–k** gave products with controllable levels of acetylation. Through adjustment of pH and appropriate selection of the glycosylation site, we were able to prepare differently acetylated glycoforms of SBL. These are illustrated by the reaction of SBL-N62C, -L217C, -S156C, and -S166C, with **1a**. Choice of pH 9.5 and SBL-N62C, -S156C, and -S166C gave tetraacetylated gluco-CMMs N62C-S-β-Glc$(Ac)_4$, S156C-S-β-Glc$(Ac)_4$, and S166C-S-β-Glc$(Ac)_4$. Alternatively, choice of SBL-L217C and pH 9.5, 7.5, or 5.5 allowed complete deacylation or incorporation of two or three acetate groups and the formation of L217C-S-β-Glc, L217C-S-β-Glc$(Ac)_2$, and L217C-S-β-Glc$(Ac)_3$, respectively. The flexibility of this method was further extended through incubation of L217C-S-β-Glc$(Ac)_3$ in pH 9.5 buffer, which allowed the clean formation of L217C-S-β-Glc in an alternative manner. In all cases, complete integrity of the site selectivity was retained. This ability to modulate the level of acetylation through pH control vastly expands the structural variety of glyco CMMs that can be conveniently accessed. Its potential in this regard was demonstrated by preparing a small library of differently acetylated glycosylated CMMs through the reaction of SBL-N62C, -S156C, -S166C, and -L217C with MTS reagents **1g–k**. For example, at pH 5.5 reactions of SBL-L217C and

Scheme 3[a]



[a] Reagents and conditions: (i) Ac_2O, py 92% for **5d**, 99% for **5e**; Ac_2O, NaOAc, 82% for **5f**; (ii) $Br(CH_2)_2OH$, $BF_3{\cdot}Et_2O$, DCM, 70% for **4i**, 67% for **4j**, 53% for **4k**; (iii) NaOMe, MeOH, 96% for **4d**, 92% for **4e**, 90% for **4f**; (iv) $NaSSO_2CH_3$, DMF, 50 °C, 80% for **1d**, 71% for **1e**, 68% for **1f**, 88% for **1i**, 83% for **1j**, 88% for **1k**.

-S166C created singly deacetylated CMMs, with the exception of pure dideacetylated glyco-CMMs L217C-S-Et-α-Glc$(Ac)_2$ and S166C-S-lac$(Ac)_5$. In contrast, reactions of SBL-S156C gave only the fully acetylated CMMs, S156C-S-**g–k**. Interestingly, the reactions of SBL-N62C were also determined by the anomeric configuration of **1g–k** with α-MTS reagents **1g,i**, giving products that retained all acetate groups while β-MTS reagents **1h,j,k** were monodeacetylated.

We attribute this valuable site-dependent deacetylation to a novel intramolecular SBL-catalyzed process. Although acetate esters are moderately chemically labile in aqueous solution at pH 9.5, they are stable at both pH 7.5 and 5.5.[18] The striking differences in behavior during modification with **1a** between SBL-L217C and the three other mutants -N62C, -S156C, and -S166C under identical reaction conditions discount the possibility of deacetylation prior to modification. Primary acetate groups are typically more labile than secondary acetate groups under conditions of enzymatic deacetylation[19] and subtilisin is regioselective for the $-CH_2OAc$ of pyranosides.[20] We believe that partial or complete deacetylations of the secondary acetates at pHs 5.5, 7.5, and 9.5 occur following sequential acetate migrations from O-4 → 6, then O-3 → O-4 → O-6, etc., that transfer, in turn, each of the secondary acetate groups to the enzyme-sensitive primary O-6 position.

Acknowledgment. We acknowledge the generous financial support of the Natural Sciences and Engineering Research Council of Canada and Genencor International, Inc., and we thank Drs. Rick Bott and Tom Graycar for helpful discussions.

Supporting Information Available: Preparation of MTS reagents, chemical modification procedures and characterization of CMMs (43 pages).

JO9816461

(18) *Protective Groups In Organic Synthesis*, 2nd ed.; Greene, T. W., Wuts, P. G. M., Eds.; Wiley: New York, 1991.
(19) Bashir, N. B.; Phythian, S. J.; Reason, A. J.; Roberts S. *J. Chem. Soc., Perkin Trans. 1* **1995**, 2203–2222.
(20) *Enzymes In Synthetic Organic Chemistry*; Wong, C.-H., Whitesides, G. M., Eds.; Pergamon/Elsevier: Oxford, 1994; pp 60–63.